

Data Analytics Task

You are tasked with building a machine-learning model to predict the closing price of a given stock based on historical data.

Duration: 5 Calendar Days

Dataset:

You are provided with 4 data sets for stock-listed companies, containing historical stock data. The dataset has the following columns:

- Date: The date of the data point
- Open: The opening price of the stock
- High: The highest price of the stock during the day
- Low: The lowest price of the stock during the day
- Close: The closing price of the stock

Requirements:

Use Python and any relevant libraries for data analysis and machine learning (e.g., Pandas, NumPy, Scikit-Learn).

Your solution should include the following:

- Data preprocessing (e.g., handling missing values, scaling features)
- Model selection and training (e.g., Linear Regression, Random Forest)
- Evaluation of the trained model (e.g., using appropriate metrics such as Mean Absolute Error, Mean Squared Error)
- Making predictions for a given set of input data
- Provide comments in your code to explain each step.

Choose any company data set to build the algorithm with and test it.

After finishing, apply the same algorithm to other companies and show the accuracy.

Submission:

- Submit your Python script GitHub link
- Brief explanation of your approach
- Any insights gained during the process.
- Present your results in a PowerPoint presentation in 15 minutes.

Hint: Try testing 80/20 and check the accuracy of your model against the four data sets. This test will assess the intern's ability to handle data, preprocess it, select and train a suitable model, and evaluate the model's performance for predicting stock prices.